Malicious entities are in vain : preserving privacy in publish and subscribe systems

Cui, S, Belguith, S, De Alwis, P, Asghar, MR and Russello, G

http://dx.doi.org/10.1109/TrustCom/BigDataSE.2018.00238

Malicious Entities are in Vain: Preserving Privacy in Publish and Subscribe Systems

Shujie Cui, Sana Belguith, Pramodya De Alwis, Muhammad Rizwan Asghar, and Giovanni Russello
Cyber Security Foundry
The University of Auckland
Auckland, New Zealand
Email: {scui379,sbel452,pdea717}@aucklanduni.ac.nz, {r.asghar,g.russello}@auckland.ac.nz

***Abstract*—Publish and subscribe (pub/sub) system is a decoupled communication paradigm that allows routing of publications. Through a set of dedicated third party servers, referred to as brokers, publications are disseminated without establishing any link between publishers and subscribers. However, the involvement of these brokers raises security and privacy issues as they can harvest sensitive data about subscribers. Furthermore, a malicious broker may collude with malicious subscribers and/or publishers to infer subscribers' interests. Our solution is such that subscribers' interests are not revealed to curious brokers and published data can only be accessed by the authorised subscribers. Moreover, the proposed protocol is secure against the collusion attacks between malicious brokers, publishers, and subscribers.**

***Index Terms*—Pub/sub, Subscribers' Privacy, Publications' Confidentiality, Collusion Resistance.**

I. Introduction

Publish and subscribe (pub/sub) system is a decoupled communication paradigm that allows data broadcasting without any link between the sender (*a.k.a. publisher*) and the receiver (*a.k.a. subscriber*). In pub/sub systems, subscribers register their interests in published data (*a.k.a. publications*) generated by publishers through a set of constraints on these data (*a.k.a. subscriptions*). Publications are routed to the interested subscribers using a network of dedicated servers referred to as *brokers*. A publication is composed of a set of tags defining a set of keywords that characterise the data content. Brokers match the publications' tags against the stored subscriptions to identify the interested subscribers. Then, the brokers filter and forward the publications to the subscribers.

Despite the benefits of pub/sub systems, they also raise security and privacy issues as the subscriptions and publications are stored and routed via dedicated brokers that could be compromised, hacked, or sniffed by adversaries [1], [2]. Indeed, publishers/subscribers may send/receive sensitive publications such as military data, health information, religious, or political interests. Thus, compromised brokers could collect sensitive information about the publishers and subscribers.

Encryption techniques are usually applied to protect sensitive information against untrusted parties in pub/sub systems. For instance, in [3]–[5], subscribers encrypt their subscriptions before registering at brokers, and publishers also encrypt publications and tags before forwarding to the brokers. Moreover, the brokers can match the subscriptions against the publications' tags on encrypted form without learning their content. However, few research works have considered the collusion attacks between malicious subscribers and brokers [6]–[8]. Indeed, a malicious broker may collude with a compromised subscriber to register her subscriptions in cleartext. Using these subscriptions, the broker can still learn information about honest subscribers' interests by checking if they match against the same publications as the compromised subscriptions.

Another limitation facing pub/sub systems is the collusion between brokers and publishers. State of the art works do not consider the publisher as an untrusted entity [9]–[11]. Specifically, a malicious publisher may collude with a compromised broker to publish a compromised publication. By identifying the subscriptions that match the compromised publication, the broker and the publisher are able to infer the subscribers' interests.

Above all, to guarantee confidentiality of both the publications and subscriptions sufficiently, a secure pub/sub system has to satisfy the following requirements:

- R1. The published data should be protected from brokers and unauthorised subscribers, *i.e.,* the publications should not be accessed by brokers and unauthorised subscribers whose interests do not match the publications' tags, even if they collude together.
- R2. The broker should be able to check if subscribers' interests match the publication tags without knowing their content, which can reveal information about the content and subscriptions.
- R3. A publisher should not be able to trace subscribers, *i.e.,* publishers and subscribers should be loosely-coupled.
- R4. The broker should not be able to know the subscribers' interests, even if it colludes with malicious subscribers or malicious publishers.

In this paper, we provide a privacy-preserving pub/sub system that meets all the requirements. Basically, to meet R1, we encrypt publications using the key policy attribute-based encryption scheme. Furthermore, we apply a Searchable Encryption (SE) scheme to enable encrypted matching between tags and subscriptions (*i.e.,* R2 and R3). The main idea to achieve R4 is to employ multiple types of brokers and divide the matching operations between encrypted subscriptions and

tags into different phases, where each phase is performed by a different type of broker. Each broker type only processes partial information from which sensitive information about encrypted interests can not be inferred. Thus, if a broker is compromised or colluding with a subscriber or a publisher, the subscriptions are still protected.

The rest of this paper is organised as follows: Section II reviews related work. We present system model, threat model, and a brief overview of our approach in Section III. Finally, we conclude this paper in Section IV.

II. Related Work

In pub/sub systems, it is crucial to protect publications' contents from unauthorised access. In addition, subscribers may want to keep their interests hidden from other subscribers as well as brokers. To deal with these issues, several research works have proposed various schemes to protect subscribers' interests against curious brokers.

In [12], Ion *et al.* present a pub/sub system that ensures confidentiality of publications and subscriptions. Their scheme allows the publishers to express fine-grained access control on the publications by applying Attribute-Based Encryption (ABE) [13] on the payload. Moreover, their scheme supports multi-user access without requiring the publishers and subscribers to share any key. However, their scheme is vulnerable to collusion attacks. That is, when the broker colludes with a malicious subscriber or publisher, they can infer the subscriptions of an honest subscriber.

In [14], Naveel *et al.* present an approach based on both symmetric and asymmetric schemes. Specifically, the publication payload is encrypted with a symmetric algorithm, and both tags and filters are encrypted with the Paillier homomorphic cryptosystem [15], such that the brokers can perform privacy-preserving matching over encrypted data. This solution offers confidentiality of publications and subscriptions. However, it breaks the de-coupling property of pub/sub system, since the subscribers have to communicate with publishers to get the subscriptions blinded. This issue has been solved in [16] by using modified Paillier cryptosystem and Attribute-Based Group Key Management (AB-GKM) scheme [17]. However, these solutions fail to prevent the broker from inferring the subscriber's interests by colluding with malicious subscribers or publishers.

Crescenzo *et al.* [18] design a 3-party pub/sub protocol that safeguards privacy of subscriptions and publications while guaranteeing performance of the system. In the protocol, both interests and tags are encrypted with 2-layer cryptographic pseudonyms, and the encrypted tags and interests are semantically secure. A trusted third party server is employed to perform the second layer of encryption. Due to the assistance of the third party, the broker is able to test the equality between encrypted tags and interests efficiently. However, in this protocol, the publication payload is encrypted with a key shared among all the subscribers and publishers, which will put all the publications at risk when the broker colludes with a malicious subscriber or publisher.

PIDGIN [19] has been proposed to ensure subscriptions' privacy and publications' confidentiality in pub/sub systems. In this proposal, the publication payload is encrypted using CP-ABE with respect to access structures. The publication tags and subscriptions are encrypted using public-key encryption with keyword search (PEKS) [20], so as to the broker could perform the matching over them without requiring access the content. However, if the broker colludes with the subscriber, the broker will be able to infer the interests of honest subscribers.

Yang *et al.* [9] introduce a dual-policy attribute-based encryption scheme that ensures an efficient keyword search in cloud-based pub/sub systems. In this proposal, the publisher defines an access policy over the publications' keywords while the subscriber sets a different access policy through its interests. In this solution, the publishers are considered fully trusted, the subscribers are malicious and the cloud server is curious. Moreover, they assume that the subscribers can collude together to access the publications but can not collude with the cloud server.

In [21], Borcea *et al.* propose PICADOR, a secure topic-based pub/sub system based on the use of a proxy-re-encryption scheme. The authors apply a lattice-based proxy re-encryption scheme that allows partial homomorphic operations. That is, the brokers have to re-encrypt the publications such that the authorised subscribers could recover the plaintext of these publications. However, this re-encryption increases the computation overhead significantly on the broker end, and the topic of each publication is sent to the broker in plaintext.

Although the aforementioned solutions ensure the publications' confidentiality, they do not consider the privacy of subscriptions against colluding brokers and subscribers [6]. In fact, a malicious subscriber can share her subscriptions in cleartext with the broker, which can leak the subscriptions of honest subscribers. This issue was addressed by Rao *et al.* in [7], [8]. Since then, all the proposals have assumed that the broker can not collude with any subscriber [9], [12], [19].

In [8] and [7], Rao *et al.* use a trusted engine to cloak the subscriptions before sending to the broker. As a result, the subscribers get more publications than they require. Although it is difficult to infer the subscribers' interests, another round of matching should be performed on the subscribers to filter out the redundant publications. Moreover, the trusted engine can be a bottleneck in the distributed pub/sub system as it must remain active and uncorrupted throughout the lifetime of the system.

More recently, Pires *et al.* [22] present a pub/sub routing engine that takes advantage of the trusted execution environment provided by shielded SGX enclaves [23]. In this approach, subscriptions are stored in the trusted SGX enclave and the match operation between interests and tags is also performed by the SGX enclave. In this case, when the brokers collude with malicious subscribers or subscribers, they can not infer other subscriptions, since the brokers can not perform the search operation. However, the subscribers have to first send the subscription for re-encryption, which violates the

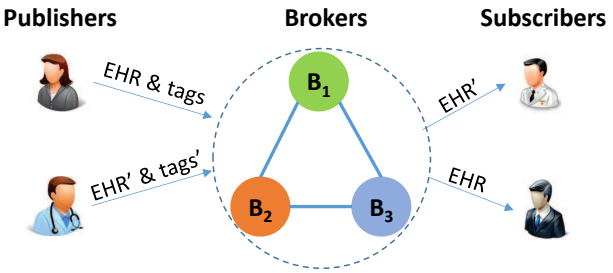

Fig. 1. An Overview of Our Proposed System: Three brokers B_1, B_2, and B_3 in different domains are connected into a virtual cluster. The publishers in these domains send publications to the cluster. The three brokers in the cluster perform the matching and routing separately, and finally only the subscribers whose interests match the tags could get the publications.

decoupling property of the pub/sub systems.

Above all, state of the art pub/sub security solutions have not considered data injection attack achieved by a compromised publisher. Indeed, a malicious publisher may generate a malicious publication and try to compromise the privacy of the interested subscribers by colluding with the broker. To do so, the malicious publisher colludes with a broker to identify the subscribers whose interests match the compromised publication. Hence, the publisher and the broker infer the subscribers' interests.

III. Solution Overview

A. Motivating Scenario

In e-health systems, medical entities (such as doctors, hospitals, clinics, and pharmacists) benefit from pub/sub services by employing private or public brokers to share patients' Electronic Health Records (EHR).

To effectively diagnose and treat patients, a publisher, say a doctor from hospital A, may need to share an EHR with other authorised doctors from hospital B, pharmacists, or a medical laboratory. In this case, the shared EHR contains personal information about the patient such as her identity, address, nature of the test, and file content. This information must be routed to various health organisations, possibly geographically separated and in independent administrative domains, where the patient can be moved when her conditions stabilise or where the tests have to be performed or analysed.

It is noteworthy that the preservation of the publication's confidentiality is not the only security concern. It is crucial to ensure confidentiality of the publication tags (including name, address of the patient, and nature of the test), representing highly sensitive information.

In addition, subscriptions are also highly sensitive information as they can reveal which patient is treated by which clinic or for which type of disease. The system should not reveal any private information related to a doctor as well as patients' EHRs. The disclosure of such information can lead to serious consequences. For example, an insurance company learning information about the health state of a patient can refuse to cover her undergoing medical tests. Basically, to provide a secure privacy-preserving pub/sub service, the system should protect the publications' confidentiality as well as the subscriptions.

B. System Model

As shown in Fig 1, we consider a privacy-preserving data pub/sub service involving the following entities:

- ***Publishers (Pub).*** The publisher generates publications and the related tags. Before publishing to the broker, she encrypts both the tags and the content of the publication.
- ***Subscribers (Sub).*** Each subscriber defines a subscription policy in the form of filters according to her interests, such that she receives only the publications whose tags satisfy the subscription policies.
- ***Broker (B).*** The broker is responsible for filtering and delivering publications to the interested Subs.
- ***Trusted Authority (TA).*** The trusted authority is responsible for managing the keys of Subs and Pubs.

C. Threat Model

In this work, we consider that the TA is fully trusted and the channels between the TA and the Pubs/Subs are secure. In our system, we consider the following threat model:

- ***Malicious Sub.*** A malicious Sub may try to access unauthorised publications and infer other Subs' interests by colluding with brokers.
- ***Malicious Pub.*** A malicious Pub may try to infer Subs' interests by injecting malicious publications and colluding with brokers.
- ***Honest but Curious Broker.*** The brokers are semi-trusted (honest-but-curious) in the system. They obey the protocol to evaluate the filters but they are curious about the content of publications and interests. Moreover, a broker may collude with any Sub or Pub to infer the other Subs' interests. In our setting, we consider that at least three brokers should be present to perform the publish services. Moreover, we assume that the malicious Sub and Pub could collude with at most two of the brokers.

D. Our Approach

In this paper, we aim at providing a pub/sub service that could protect publications and Subs' interests from curious brokers in the presence of malicious Subs and Pubs.

To achieve R1, *i.e.,* to protect the publications from unauthorised entities, the Pub can encrypt the publication using Key-Policy Attribute-Based Encryption (KP-ABE) scheme [24]. On the one hand, the confidentiality of the publication can be protected. On the other hand, the Pub could control the access over her publications by defining the access control structure. For achieving R2, tags and interests could be encrypted using an SE scheme. Thus, the brokers could check if the publication tags match Subs' interests in an encrypted manner, and distribute the publication to authorised Subs (*i.e.,* R3).

Encrypting Sub interests using SE is not sufficient to achieve R4. As mentioned above, when the broker colludes with malicious Pubs or Subs, it can infer the honest Subs' interests by observing the matching results. The novelty of our proposal

lies in the fact that Subs' interests are kept protected even when a broker colludes with a malicious Sub or Pub. Unlike state-of-the-art pub/sub systems that fundamentally use a single broker to match and forward the publications to the Subs, our solution is based on the use of three different types of brokers. The main idea of this proposal is to divide the matching operations between interests and tags into three different phases where each phase is performed by a different type of broker. Basically, the Sub defines her filter as a tree whose leaves represent interests and non-leaf nodes denote AND, OR and NOT gates. The leaves and non-leaf nodes are sent to two different brokers separately. Furthermore, the leaves are encrypted with SE and permuted with a keyed Pseudo-Random Permutation (PRP) before sending to the broker, and the key of the PRP is sent to the third broker. The broker who gets the interests is responsible for matching each interest against the corresponding publication tag in encrypted form. The broker who gets the key of PRP will recover the order of the matching results by inverting the permutation. The third broker evaluates the tree and generates the final matching result. If the Sub's interests match the publication's tags, the third broker forwards the publication to the Sub.

In our solution, each type of brokers only knows some partial information, from which sensitive information about encrypted interests can not be inferred. Thus, if a malicious Sub or Pub colludes with one or two types of the brokers, they are unable to infer the interests of honest Subs.

IV. Conclusions and Future Work

In pub/sub systems, publications are disseminated to interested subscribers through a set of untrusted brokers. These brokers may collect sensitive information by accessing publication tags and subscribers' interests. In addition, a malicious broker can collude with compromised publisher and/or subscribers to infer subscribers' interests. To mitigate this issue, we introduce a novel design of pub/sub systems to protect the subscribers' interests against curious brokers. Moreover, the proposed solution is resistant against the collusion attacks between a broker and a subscriber.

As future work, we aim to introduce the details of proposed pub/sub system. In addition, we aim to implement a prototype to show the feasibility and efficiency of our solution.

Acknowledgements

This research is supported by STRATUS (Security Technologies Returning Accountability, Trust and User-Centric Services in the Cloud), a project funded by the Ministry of Business, Innovation and Employment (MBIE), New Zealand.

References

[1] C. Esposito, M. Ciampi, and G. De Pietro, "An event-based notification approach for the delivery of patient medical information," *Information Systems*, vol. 39, pp. 22–44, 2014.

[2] M. Cinque, C. Di Martino, and C. Esposito, "On data dissemination for large-scale complex critical infrastructures," *Computer Networks*, vol. 56, no. 4, pp. 1215–1235, 2012.

[3] C. Esposito and M. Ciampi, "On security in publish/subscribe services: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 966–997, 2015.

[4] B. Shand, P. Pietzuch, I. Papagiannis, K. Moody, M. Migliavacca, D. Eyers, and J. Bacon, "Security policy and information sharing in distributed event-based systems," *Reasoning in Event-Based Distributed Systems*, pp. 151–172, 2011.

[5] G. Di Crescenzo, J. Burns, B. Coan, J. Schultz, J. Stanton, S. Tsang, and R. N. Wright, "Efficient and private three-party publish/subscribe," in *International Conference on Network and System Security*. Springer, 2013, pp. 278–292.

[6] E. Onica, P. Felber, H. Mercier, and E. Rivière, "Confidentiality-preserving publish/subscribe: A survey," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, p. 27, 2016.

[7] W. Rao, L. Chen, M. Yuan, S. Tarkoma, and H. Mei, "Subscription privacy protection in topic-based pub/sub," in *International Conference on Database Systems for Advanced Applications*. Springer, 2013, pp. 361–376.

[8] W. Rao, L. Chen, and S. Tarkoma, "Toward efficient filter privacy-aware content-based pub/sub systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 11, pp. 2644–2657, 2013.

[9] K. Yang, K. Zhang, X. Jia, M. A. Hasan, and X. S. Shen, "Privacy-preserving attribute-keyword based data publish-subscribe service on cloud platforms," *Information Sciences*, vol. 387, pp. 116–131, 2017.

[10] E. Onica, P. Felber, H. Mercier, and E. Rivière, "Efficient key updates through subscription re-encryption for privacy-preserving publish/subscribe," in *Proceedings of the 16th Annual Middleware Conference*. ACM, 2015, pp. 25–36.

[11] Y. Polyakov, K. Rohloff, G. Sahu, and V. Vaikuntanthan, "Fast proxy re-encryption for publish/subscribe systems," *IACR Cryptology ePrint Archive*, vol. 2017, p. 410, 2017.

[12] M. Ion, G. Russello, and B. Crispo, "Design and implementation of a confidentiality and access control solution for publish/subscribe systems," *Computer networks*, vol. 56, no. 7, pp. 2014–2037, 2012.

[13] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *S&P 2007*. IEEE Computer Society, 2007, pp. 321–334.

[14] M. Nabeel, N. Shang, and E. Bertino, "Efficient privacy preserving content based publish subscribe systems," in *Proceedings of the 17th ACM symposium on Access Control Models and Technologies*. ACM, 2012, pp. 133–144.

[15] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *EUROCRYPT 1999*, ser. Lecture Notes in Computer Science, J. Stern, Ed., vol. 1592. Springer, 1999, pp. 223–238.

[16] M. Nabeel, S. Appel, E. Bertino, and A. Buchmann, "Privacy preserving context aware publish subscribe systems," in *International Conference on Network and System Security*. Springer, 2013, pp. 465–478.

[17] M. Nabeel and E. Bertino, "Attribute based group key management," *Trans. Data Privacy*, vol. 7, no. 3, pp. 309–336, 2014.

[18] G. D. Crescenzo, J. Burns, B. A. Coan, J. L. Schultz, J. R. Stanton, S. Tsang, and R. N. Wright, "Efficient and private three-party publish/subscribe," in *NSS 2013*, ser. Lecture Notes in Computer Science. Springer, 2013, pp. 278–292.

[19] M. R. Asghar, A. Gehani, B. Crispo, and G. Russello, "PIDGIN: Privacy-preserving interest and content sharing in opportunistic networks," in *Proceedings of the 9th ACM symposium on information, computer and communications security*. ACM, 2014, pp. 135–146.

[20] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *EUROCRYPT 2004*, ser. Lecture Notes in Computer Science, C. Cachin and J. Camenisch, Eds., vol. 3027. Springer, 2004, pp. 506–522.

[21] C. Borcea, Y. Polyakov, K. Rohloff, G. Ryan *et al.*, "Picador: End-to-end encrypted publish–subscribe information distribution with proxy re-encryption," *Future Generation Computer Systems*, vol. 71, pp. 177–191, 2017.

[22] R. Pires, M. Pasin, P. Felber, and C. Fetzer, "Secure content-based routing using intel software guard extensions," in *Middleware 2016*. ACM, 2016, p. 10.

[23] V. Costan and S. Devadas, "Intel SGX explained," *IACR Cryptology ePrint Archive*, vol. 2016, p. 86, 2016.

[24] R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *CCS 2007*. ACM, 2007, pp. 195–203.